IOP Publishing Environ. Res. Lett. 17 (2022) 055002 https://doi.org/10.1088/1748-9326/ac6229

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
10 December 2021

REVISED
16 March 2022

ACCEPTED FOR PUBLICATION
29 March 2022

PUBLISHED
19 April 2022

LETTER

Regionalization of climate teleconnections across Central Asian mountains improves the predictability of seasonal precipitation

Atabek Umirbekov[1,2,3,*], **Mayra Daniela Peña-Guerrero**[1,2,3] and **Daniel Müller**[1,2,4]

[1] Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany
[2] Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
[3] Tashkent Institute of Irrigation and Agricultural Mechanization Engineers (TIIAME), 39 Kari Niyazov Str., Tashkent 100000, Uzbekistan
[4] Integrative Research Institute on Transformations of Human-Environment Systems (IRI THESys), Humboldt Universität-zu-Berlin, Berlin, Germany
[*] Author to whom any correspondence should be addressed.

E-mail: umirbekov@iamo.de

Keywords: climate teleconnections, Central Asia, machine learning, mountains, seasonal forecasting, precipitation

Supplementary material for this article is available online

Abstract

Mountains play a critical role in water cycles in semiarid regions by providing for the majority of the total runoff. However, hydroclimatic conditions in mountainous regions vary considerably in space and time, with high interannual fluctuations driven by large-scale climate oscillations. Here, we investigated teleconnections between global climate oscillations and the peak precipitation season from February to June in the Tian-Shan and Pamir Mountains of Central Asia. Using hierarchical climate regionalization, we identified seven subregions with distinct precipitation patterns, and assessed correlations with selected climate oscillations at different time lags. We then simulated the seasonal precipitation in each subregion from 1979 to 2020 using the most prevalent teleconnections as predictors with support vector regression (SVR). Our findings indicate that the El Niño–Southern Oscillation, the Pacific Decadal Oscillation, and the Eastern Atlantic/West Russia pattern are among the major determinants of the seasonal precipitation. The dominant lead-lag times of these oscillations make them reliable predictors ahead of the season. We detected notable teleconnections with the North Atlantic Oscillation and Scandinavian Pattern, with their strongest associations emerging after onset of the season. While the SVR-based models exhibit robust prediction skills, they tend to underestimate precipitation in extremely wet seasons. Overall, our study highlights the value of appropriate spatial and temporal aggregations for exploring the impacts of climate teleconnections on precipitation in complex terrains.

1. Introduction

Mountains play a crucial role in providing water resources, particularly in arid and semiarid regions, where they act as 'water towers' by providing the majority of the total runoff (Viviroli and Weingartner 2004). However, large-scale ocean-atmosphere oscillations can cause high interannual hydroclimatic variability, particularly in semiarid regions (Scholes 2020). In regions where livelihoods depend on river runoff for irrigating agricultural land, mountain droughts that reduce downstream runoff can be a particular bane. Early warning systems that provide accurate seasonal precipitation forecasts can permit better preparation for water shortages (Portele *et al* 2021).

Landlocked in the heart of the Asian continent, the Pamir and Tian-Shan Mountains are vital for water cycles in Central Asia, since these mountains provide the largest fraction of river discharge (Immerzeel *et al* 2020). High interannual hydroclimatic fluctuations in these mountain ranges impose substantial burdens on irrigated crop production that dominates the southern parts of Central Asia (Karthe *et al* 2017). Better preparing for variable downstream river runoff has long been called for, such as through implementing early warning systems based on seasonal forecasts (Karthe *et al* 2017,

Apel *et al* 2018, World Bank 2018, Gerlitz *et al* 2019, Xenarios *et al* 2019). In this regard, global and regional climate oscillations have been suggested as viable indicators for long-range seasonal hydrological outlooks in the region (Gerlitz *et al* 2020).

Much of the previous research on the relationship between climate oscillations and precipitation variability focused on entire Central Asia and explored the effects of the El Niño-Southern Oscillation (ENSO). It was reported that ENSO during its El Niño phase increases precipitation intensity across Central Asia, with the amplitude of this effect being particularly strong from the autumn to summer (Mariotti 2007, Chen *et al* 2018, Gerlitz *et al* 2019). In contrast, La Niña state is associated with below-average precipitation in the region. ENSO impacts can be amplified by the Pacific Decadal Oscillation (PDO): More severe droughts in Central Asia tend to occur during ENSO's La Niña phase when the PDO is in the cold phase (Wang *et al* 2014).

Recurring large ocean-atmospheric fluctuations in the Atlantic Ocean are another source of seasonal and interannual precipitation variations in Central Asia. The most prominent are the North Atlantic Oscillation (NAO), Scandinavian pattern (SCAN), and East Atlantic/Western Russia pattern (EAWR), all of which refer to periodic fluctuations in atmospheric pressure between specific regions of the Atlantic Ocean and Eurasia. These teleconnections manifest as wave trains that modulate the location and strengths of the westerlies (Bothe *et al* 2012). The Atlantic Multidecadal Oscillation (AMO) is a periodic variation in currents in the northern Atlantic Ocean, and its positive state in the winter is associated with higher precipitation in Central Asia (Gerlitz *et al* 2019).

Knowledge about the impacts of the Indian Ocean Dipole (DMI), an irregular oscillation of sea surface temperature between the western and eastern parts of the Indian ocean (Saji *et al* 1999), is patchy. Existing evidence suggests that summer runoff fluctuations in main tributaries of the Amu-Darya, the largest river in the region which originates in the Pamir mountains, are negatively associated with DMI (Dixon and Wilby 2019). However, the timing of this relationship remains unknown. Recent findings indicate a negative but weak correlation between summer precipitation in the Tian-Shan and concurrent DMI, and that it inverts and weakens over the winter season (Guan *et al* 2022).

Cyclical climatic phenomena that occur inland may also impact the precipitation variability in the region. For example, the Siberian High (SH), a high atmospheric pressure anomaly that forms across central Siberia during the winter, dominates climate variability in the Northern Hemisphere during the winter months (Cohen *et al* 2001). The SH can impact winter precipitation variability across the mountainous region of South Asia that neighbours the Pamir (Riaz and Iqbal 2017). Finally, the quasi-Biennial oscillation (QBO), which regulates the direction of stratospheric winds in the equatorial zones, is positively associated with precipitation in some areas of Central Asia, but this relationship appears statistically insignificant across the larger part of Central Asia (Brönnimann *et al* 2016, Gerlitz *et al* 2019).

Unfortunately, knowledge of how these climate oscillations affect local precipitation patterns in the mountains of Central Asia is incomplete. Using predominantly coarser gridded precipitation datasets, previous research largely focused on the larger geographical domain of Central Asia, of which the mountains of Central Asia only make up a tiny fraction. However, the complex terrain and orographic effects in the mountains drive distinct precipitation patterns, typically with high spatial heterogeneity. In addition, climate teleconnections are known to exhibit non-stationary behaviour in space and time. For instance, the relationship between SCAN and cold season precipitation in Central Asia may inverse depending on the location (Gerlitz *et al* 2019). In the mountain areas, a positive relationship between SCAN and precipitation during the winter can become negative during the summer (Schiemann 2007). These regional specificities call for a rigorous characterization of the spatial bounds of the targeted precipitation season, something that can be accomplished through climate regionalization (Badr *et al* 2016, Satti *et al* 2017).

In addition, climate teleconnections usually exhibit nonlinear behaviour, which limits the predictive power of traditional linear techniques for seasonal climate forecasting (NRC 2010, Bothe *et al* 2012). On the other hand, machine learning-based approaches reportedly can better account for nonlinearities and interactions and thus outperform traditional statistical methods (Chantry *et al* 2021, Gibson *et al* 2021). Many machine learning methods generally rely on large sample sizes, which impedes their usage in data-limited regions such as Central Asia. However, kernel-based approaches, such as support vector machines, offer more advantages in handling multivariate data in smaller samples and avoid overfitting (NRC 2010, Raghavendra and Deka 2014).

Here we have three main objectives. First, we aim to identify distinct precipitation subregions in the Tian-Shan and Pamir Mountains during the peak season via hierarchical climate regionalization. We define the peak precipitation season as the period from February to June, when precipitation and its interannual variation have the highest magnitude across most of the study area[5]. Second, we aim to determine the relationships between selected global climate oscillations and the variability of the peak

[5] Another rationale for targeting the season with high interannual variability of precipitation is to reduce uncertainties in seasonal hydrological forecasts, see supplementary material S1 to this manuscript.

season precipitation in each identified subregion. Finally, we assess the predictability of peak season precipitation with support vector regression (SVR), which uses the leading spatiotemporal modes of teleconnections as predictors.

2. Material and methods

2.1. Study area

This study focuses on the Tian-Shan Mountains that are in the territories of Kyrgyzstan, Kazakhstan, Uzbekistan, and the Pamir Mountains located in the territories of Tajikistan and Afghanistan (figure 1). The Pamir Mountains serve as the primary watershed for Amu-Darya, and the Tian-Shan Mountains serve as the primary watershed for Syr-Darya, Central Asia's two largest rivers (Schär *et al* 2004, Immerzeel *et al* 2020).

2.2. Data

We used monthly precipitation estimates from the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset, which is available at a spatial resolution of 0.1° from 1979 to 2020 (Beck *et al* 2019). MSWEP version 2.8 uses improved weight maps (GloH2O 2021) to merge precipitation estimates from gauging stations, reanalysis product ERA5 (Hersbach *et al* 2020) and satellite-based precipitation estimates from IMERG GPM (Huffman *et al* 2019). A recent comparison of the MSWEP v2.8 to independent gauging station estimates over the Pamir and Tian-Shan Mountains revealed high agreement, especially for spring precipitation (Peña-Guerrero *et al* 2022).

We used monthly indices of the climate oscillations for the period from 1978 to 2020 (table 1). The Southern Oscillation Index (SOI) here serves as a proxy of the ENSO state.

2.3. Methods

2.3.1. Regionalization of seasonal precipitation

We applied Ward's hierarchical minimum variance method (Ward 1963) to identify spatial clusters with similar precipitation patterns. It is an agglomerative hierarchical clustering technique, where the criterion for choosing the pair of clusters to merge at each step is based on minimizing the increase in the sum of squared errors. We conducted precipitation regionalization using pixelwise monthly precipitation from February to June from 1979 to 2020.

2.3.2. Associations between climate oscillations and seasonal precipitation

We calculated the Spearman's rank correlation coefficient between the seasonal precipitation averages for each subregion and each climate oscillation index at varied lead-lag times. Because the associations between oscillation indices and precipitation might change over time, we explored possible associations of each oscillation index starting from January of the preceding year until the end of June, i.e. the final month of the targeted peak season. We evaluated the field significance of the most prevalent temporal patterns on pixel-level within each subregion using the false discovery rate (FDR) approach. The FDR identifies locally significant tests by imposing more stringent constraints to the expected proportion of falsely rejected local null hypotheses (Benjamini and Hochberg 1995, Wilks 2006).

2.3.3. Development of forecast models

We used the climate oscillations with the strongest temporal lead-lags that passed the FDR test as predictors in the SVR. The seasonal precipitation series were split into training (80%) and validation (20%) samples by applying stratified partitioning based on the probability distribution function. We used the mean absolute percentage error (MAPE), squared coefficient of correlation (r^2) and Kling-Gupta efficiency (KGE) to assess the overall performance of the SVR models. Since some oscillations may have the largest impact after the start of the precipitation season, we constructed both forecast and in-season SVR models for each subregion, with the latter utilizing in-season values of those oscillations.

Note: Supplementary material S1 (available online at stacks.iop.org/ERL/17/055002/mmedia) provides an expanded description of the methodological steps summarized above.

3. Results

3.1. Spatial variability of the annual and seasonal precipitation

We identified seven subregions with distinct precipitation patterns (figure 2). The Tian-Shan Mountains separate into three precipitation subregions, which we refer to as the Western Tian-Shan, Northern Tian-Shan, and Inner Tian-Shan. In addition, we obtained a precipitation subregion in the central part of the Tian-Shan that we label Ferghana-Alay. The Pamir Mountains are also delineated into three subregions, which we denote as the Western Pamir, Central Pamir, and Eastern Pamir.

Long-term monthly precipitation within the subregions corroborates their unique precipitation patterns (figure 3). Precipitation peaks and the largest interannual changes occur during the spring months in all subregions, except in the Inner Tian-Shan and Eastern Pamir, where precipitation is skewed towards the summer season. The Western Tian-Shan and Western Pamir have the highest intra- and interannual variations. Precipitation in the Inner Tian-Shan as well as the Central and Eastern Pamir follows a less dynamic annual cycle with a relatively lower incidence of extreme deviations from long term means.

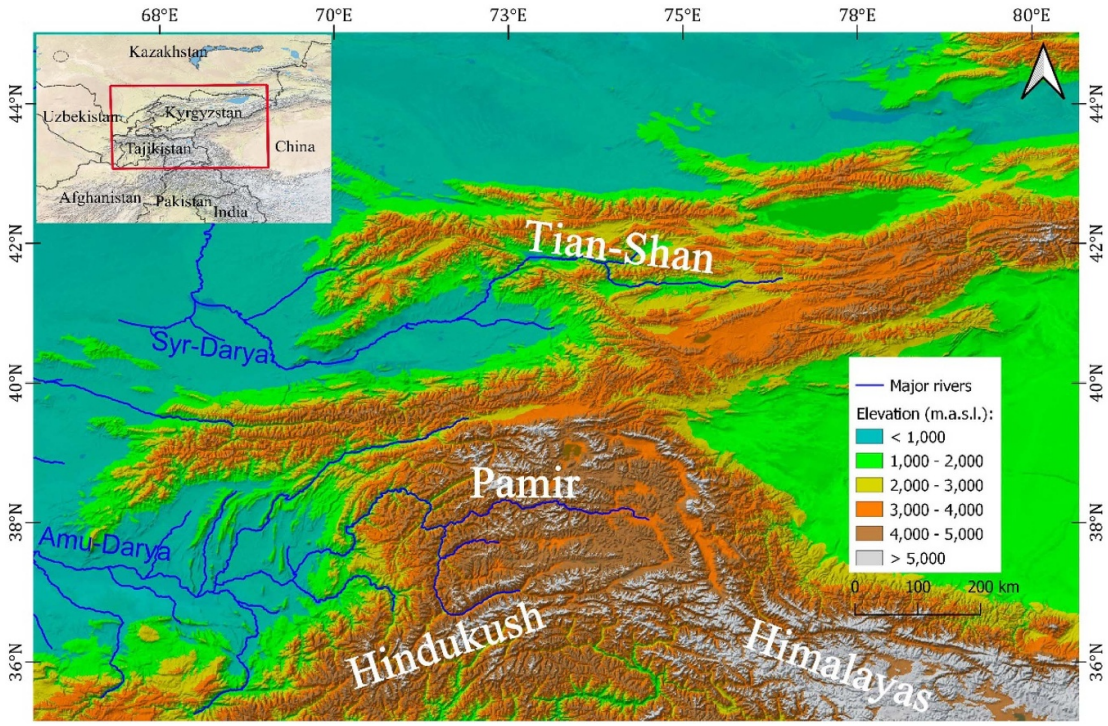

Figure 1. Location of the Tian-Shan and Pamir Mountains.

Table 1. Climate oscillation indices used in this study.

Climate oscillation index	Name	Type	Source
SOI	Southern Oscillation Index	Sea level pressure difference	Ropelewski and Jones (1987)
QBO	Quasi-Biennial Oscillation	Stratospheric zonal wind	Naujokat (1986)
PDO	Pacific Decadal Oscillation	Sea surface temperature difference	Mantua *et al* (1997)
EAWR	East Atlantic/West Russia pattern (EAWR)	Sea level pressure difference	Barnston and Livezey (1987)
NAO	North Atlantic Oscillation (NAO)	Sea level pressure difference	Barnston and Livezey (1987)
SCAN	Scandinavian pattern	Sea level pressure difference	Barnston and Livezey (1987)
AMO	Atlantic Multidecadal Oscillation	Sea surface temperature difference	Enfield *et al* (2001)
DMI	Dipole Mode Index	Sea surface temperature difference	Saji and Yamagata (2003)
SH	Siberian High	Sea level pressure difference	Reconstructed from the NCEP-NCAR reanalysis, using monthly anomalies over 40–60° N and 80–100° E

Annual and seasonal precipitation have a stronger link to geographic location than to the average elevation of a subregion (figure 3 and supplementary material S2). The Western Tian-Shan and Pamir receive the highest amounts of annual precipitation, while the eastern parts of the study area receive the least, despite having a higher elevated terrain. This disparity becomes even more pronounced when February-to-June season precipitation is considered: seasonal totals in the high-elevation Inner Tian-Shan, Central Pamir, and Eastern Pamir are barely half of the precipitation levels in the western subregions.

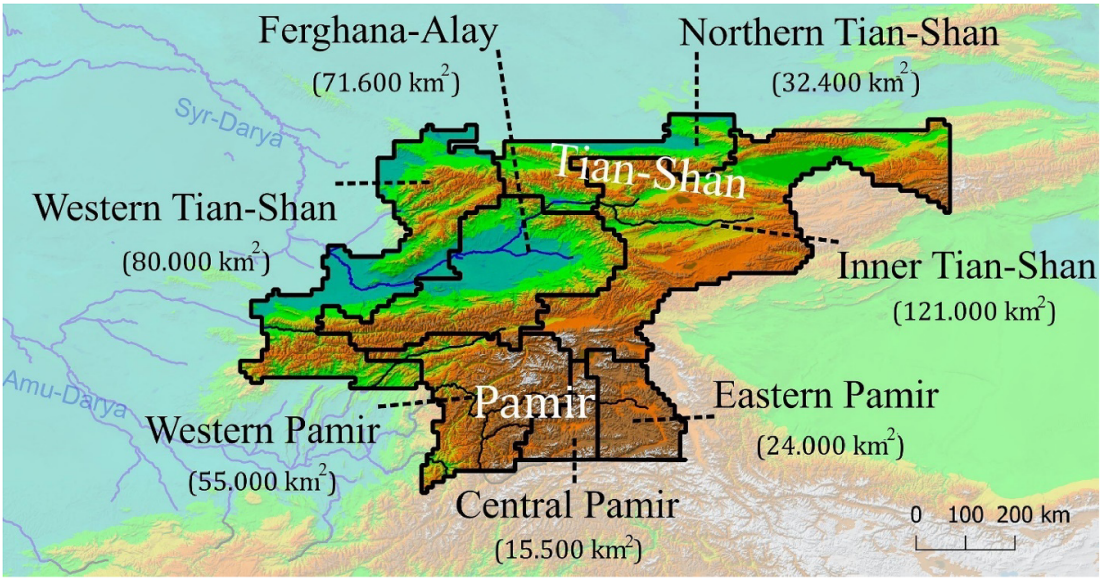

Figure 2. Spatial distribution of the precipitation subregions and their approximate areas. The background colour legend is identical to that in figure 1.

3.2. Associations between climate oscillations and seasonal precipitation

The correlations between global oscillation indices and seasonal precipitation in the study area reveal contrasting temporal patterns (figure 4). Their propagation across the subregions is, however, less heterogeneous, i.e. oscillation indices at any given month tend to have the same trend (sign) of correlation with the peak season precipitation across all subregions, except for the Eastern Pamir.

The average seasonal precipitation across most subregions is strongly correlated with SOI and PDO over a longer period than with any other oscillation. The seasonal precipitation and SOI exhibit a statistically significant negative correlation already three months prior to the start of the season in most subregions, and persists longer in the western regions. The PDO has a positive relationship with seasonal precipitation, which appears as early as seven months before the start of the season, reaching statistically significant levels by September to November.

The SCAN and NAO have a relatively shorter temporal association with the seasonal precipitation, which is strongest in the spring after the onset of the peak season. The state of SCAN in March to May shows the higher correlation, particularly for the precipitation in the elevated western and central Pamir, followed by the Inner Tian-Shan. In February and March, the NAO has a negative relationship with the seasonal precipitation in all subregions, except the Eastern Pamir. The statistically significant correlations of NAO (0.05 confidence level) however appear only in Ferghana-Alay and Inner Tian-Shan.

The EAWR and SH have stronger associations with seasonal precipitation during specific months. October EAWR has a substantial and positive correlation with seasonal precipitation over most of the study area, but is less robust in Eastern Pamir. SH in December exhibits a mild negative association that also spans all subregions, but is statistically significant at 0.05 confidence level only in Ferghana-Alay and Inner Tian-Shan.

We detected even longer lead times for the AMO and DMI, both negatively correlated with seasonal precipitation. A one year lead time of AMO and DMI results in a statistically significant correlation for precipitation over the Western and Central Pamir. The AMO extends its statistically significant correlation with the seasonal precipitation into the Inner Tian-Shan subregion. The QBO and precipitation are positively correlated for the central and northeastern parts of the study area, albeit weaker and with a longer temporal autocorrelation.

The local correlation patterns during the months when climate oscillations exhibit the strongest average association with seasonal precipitation (hereinafter referred to as 'dominant lead-lag time') show higher spatial homogeneity across subregions (figure 5). Because most associations either persist or even become stronger at longer temporal scales, the majority of the climatic oscillation indices were aggregated as 3 month averages. This aggregation did not include EAWR and SH, which have strong correlation during specific months. While SOI also demonstrated a persistently strong correlation over a longer time scale, the figure 5 depicts only correlation of SOI in December.

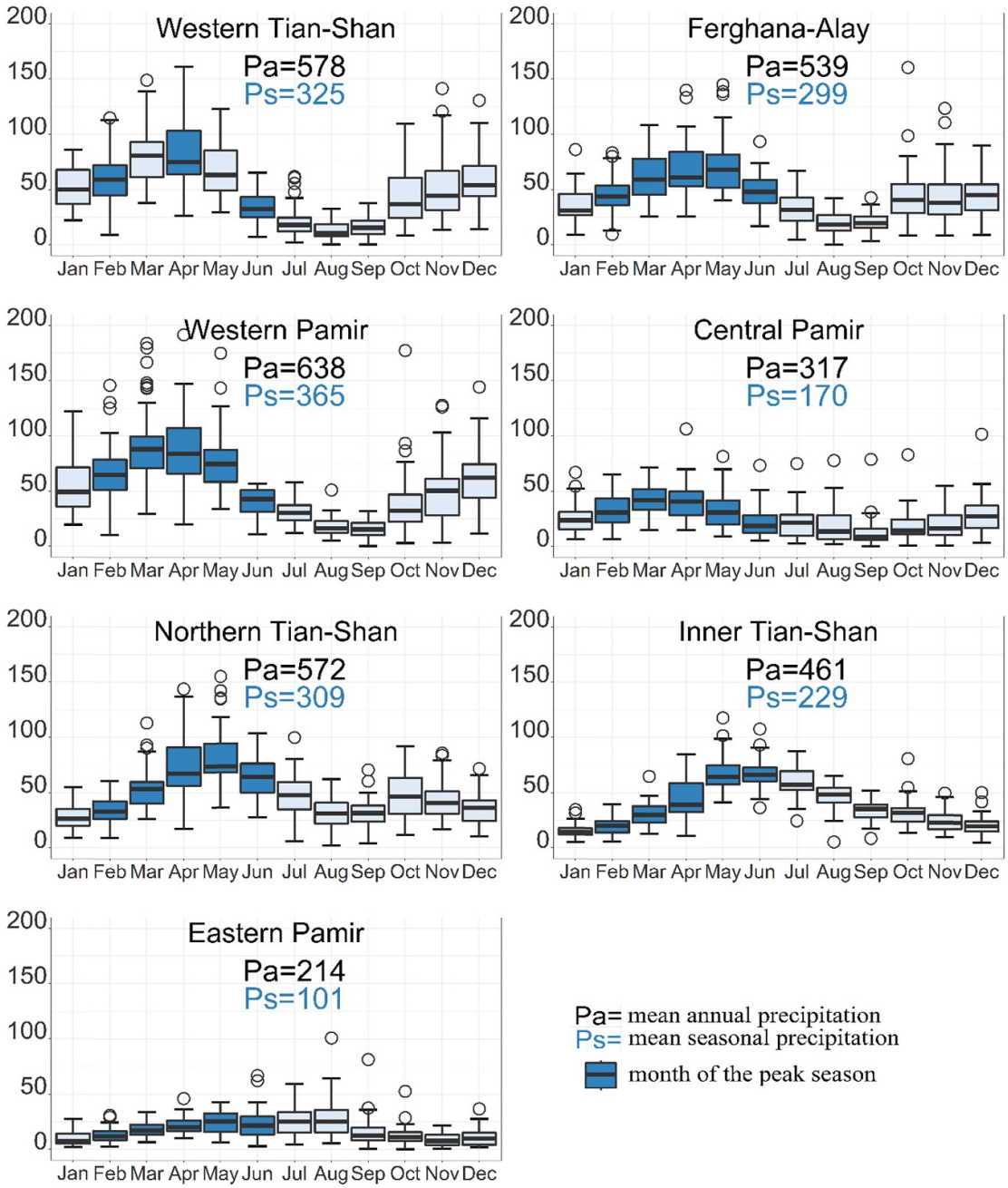

Figure 3. Long-term monthly precipitation patterns. The boxplots show the interquartile range as boxes, median as line inside the boxes, and outliers as points. All units are in millimetres.

The field significance test based on the false discovery rate (FDR) approach (table 2) leaves out approximately half of the revealed associations depicted in figure 5. It is suggested that when data is spatially autocorrelated, as climatic variables do, the control level of FDR (α_{FDR}) could be accommodated as double of a global null hypothesis test threshold (Wilks 2016). At control level of $\alpha_{FDR} = 0.1$, the field significance test rejects the null hypothesis for SOI, PDO, EAWR, and SCAN across all subregions except Eastern Pamir. The dominant lead-lag times of these four climate oscillations persist statistically significant even under a more conservative level of level of $\alpha_{FDR} = 0.05$ (see supplementary material S4). In addition, the FDR test at $\alpha_{FDR} = 0.1$ confirms statistical significance of NAO's connections with the seasonal precipitation in Inner Tian-Shan and West Pamir. AMO and DMI's correlations also meet the field significance test for Central Pamir.

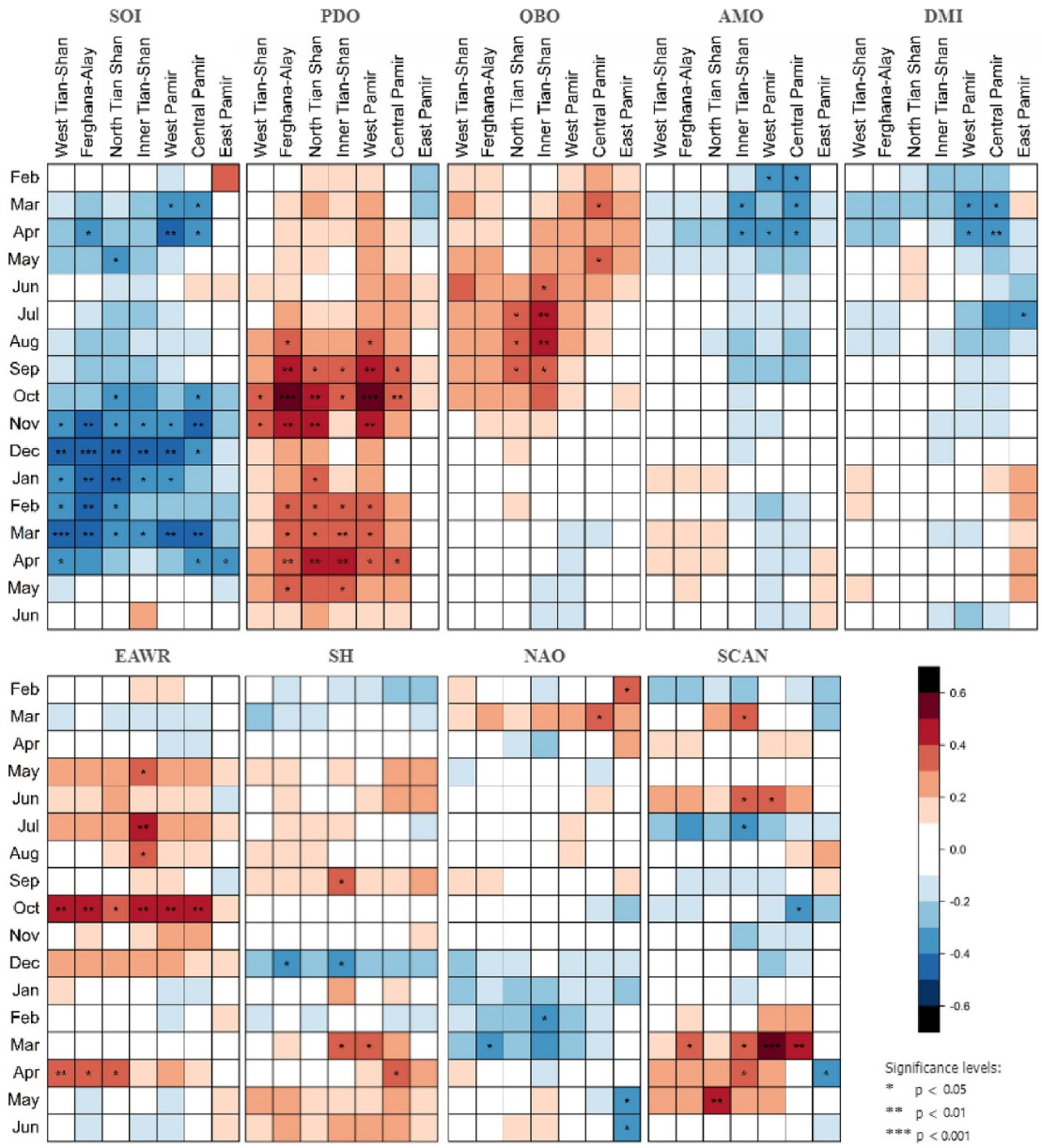

Figure 4. Spearman's rank correlation coefficients (ρ) between the climate oscillation indices and February–June precipitation for all subregions. Columns indicate the subregions; rows denote months of a climate index, with bottom February–June months concurrent with the precipitation season. Abbreviations are spelled out in table 1.

3.3. Predictability of area-averaged seasonal precipitation

Based on the field significance test results, we selected SOI, PDO, EAWR, NAO, and SCAN as candidate predictors for the SVRs (table 3). The overlapping dominant lag times of SCAN and NAO with the targeted precipitation season makes them less useful for the practical forecasting. Hence, the *forecast* SVR models omitted the former and instead used only December values of the NAO index, at which point it begins to display association across most subregions (albeit being yet statistically non-significant). To assess the influence of the selected oscillations regardless of whether they coincide with the target season, we additionally elaborated *in-season* SVR models that incorporate all five climate oscillations.

In general, the forecast SVR models show robust prediction skills for all subregions, except the Eastern Pamir where performance of r^2 and KGE were close to 0 (table 4). The mean absolute percentage error (MAPE) of the SVR forecasts was on average 15%, ranging between 7% in the Inner Tian-Shan to 16% in the Western Tian-Shan (table 4). Higher MAPE values do not necessarily imply poorer prediction skill since the forecast model may show better performance in terms of r^2 and KGE coefficients. When compared to the hindcasts produced by the

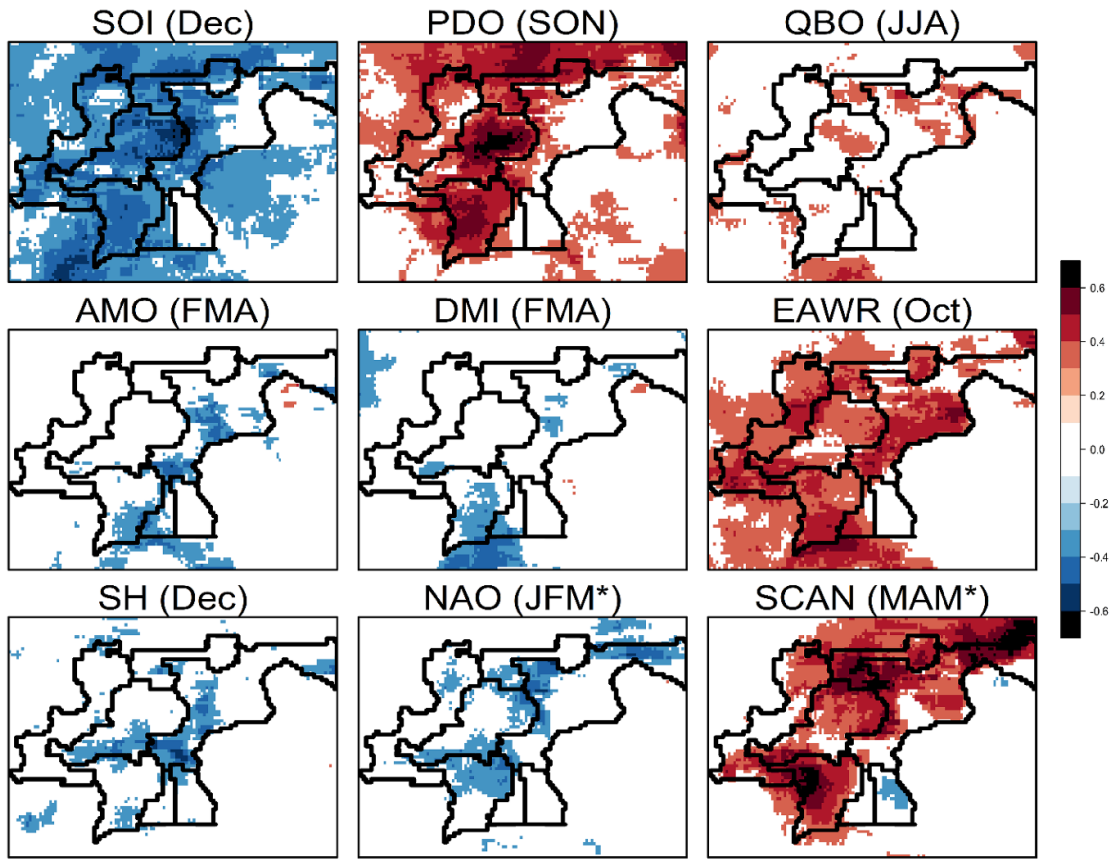

Figure 5. Spearman's rank correlation (p-value < 0.05) between climate oscillations at their dominant lead months and pixelwise precipitation from February to June. Since the oscillations of some teleconnections propagate for several consecutive months, they are indicated with their 3 month averages (e.g. PDO (SON) represents the correlation coefficients of the PDO index aggregated over September, October, and November). Superscript * denotes months of climate oscillations that coincide with the precipitation season. Abbreviations are spelled out in table 1.

Table 2. Proportion of local correlations in each subregion that passed field significance test under $\alpha_{FDR} = 0.1$. Red highlights indicate months of the climate indices concurrent with the precipitation season (February–June). For example, 'PDO (SON)' refers to the Pacific Decadal Oscillation index averaged over preceding September, October, November, whereas 'SCAN (*MAM*)' denotes the Scandinavian Pattern averaged over March, April, May months. Abbreviations are spelled out in table 1.

	SOI (Dec)	PDO (SON)	QBO (JJA)	AMO (FMA)	DMI (FMA)	EAWR (FMA)	SH (Dec)	NAO (*JFM*)	SCAN (*MAM*)
West Tian-Shan	97	83	0	0	0	99	0	0	59
Ferghana-Alay	99	95	0	0	0	97	0	0	93
North Tian-Shan	95	97	0	0	0	95	0	0	100
Inner Tian-Shan	68	27	0	19	0	74	14	28	74
West Pamir	100	100	0	0	0	100	0	55	97
Central Pamir	87	48	0	72	35	81	0	0	0
East Pamir	0	0	0	0	0		0	0	0

European Centre for Medium-Range Weather Forecasts' (ECMWF) SEAS5 forecasting system, the SVR models recreate seasonal precipitation more consistently throughout all subregions (see supplementary material S5). While the forecast SVR models simulate well the interannual variability of seasonal precipitation (figure 6), they tend to underestimate extremely wet seasons.

The in-season SVR models generally outperform their forecast counterparts, as evidenced by narrower MAPE and higher r^2 and KGE. A likely reason is that they utilize the NAO and SCAN indices for early spring months, when these indices show their strongest association with seasonal precipitation (figure 4). The incremental improvements in model accuracy are especially significant for the central part of the Pamir. The Ferghana-Alay and Northern Tian-Shan also benefit from a comparable improvement in *in-season* SVR model performance in terms of the reduced MAPE and higher r^2 values.

Table 3. Dominant lead-lag times of the climate indices used as predictors in the forecast and in-season SVR models. Red highlights indicate months of the climate indices concurrent with the precipitation season (February–June).

	Subregion	SOI	PDO	EAWR	NAO	SCAN
Forecast SVR	All subregions (excluding East Pamir)	Dec	SON	Oct	Dec	—
In-season SVR	West Tian-Shan	Dec	SON	Oct	Mar	MAM
	Ferghana-Alay	Dec	SON	Oct	Mar	MAM
	North Tian-Shan	Dec	SON	Oct	JFM	MAM
	Inner Tian-Shan	Dec	SON	Oct	JFM	MAM
	West Pamir	Dec	SON	Oct	JFM	MAM
	Central Pamir	Dec	SON	Oct	JFM	Mar
	East Pamir	JFM			MAM	MAM

Table 4. Performance of the resulting forecast and in-season SVR models on the validation set.

	Forecast SVR			In-season SVR		
Subregion	MAPE	r^2	KGE	MAPE	r^2	KGE
West Tian-Shan	16%	0.40	0.44	16%	0.57	0.41
Ferghana-Alay	14%	0.53	0.65	12%	0.68	0.48
Northern Tian-Shan	13%	0.41	0.36	10%	0.72	0.55
Inner Tian-Shan	7%	0.47	0.52	7%	0.46	0.55
Western Pamir	15%	0.42	0.25	15%	0.55	0.29
Central Pamir	15%	0.37	0.34	14%	0.64	0.30
Eastern Pamir	—	—	—	20%	0.16	0.20

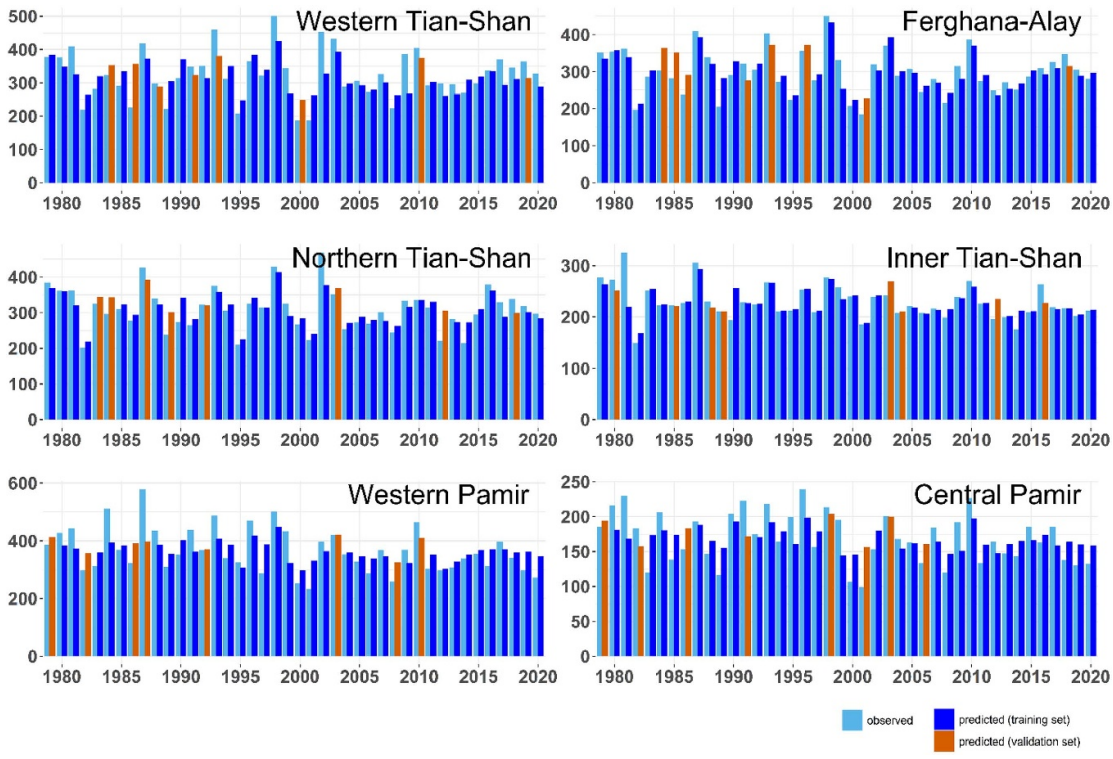

Figure 6. Comparison of the observed seasonal precipitation totals (mm) and precipitation predicted by the forecast SVR models. Note: The scale of precipitation (y-axis) differs across each subregion graph.

4. Discussion and conclusions

The Pamir and Tian-Shan Mountains of Central Asia have highly heterogeneous precipitation patterns, with their western areas receiving significantly more precipitation and experiencing greater interannual variability than their eastern ranges. Most annual precipitation in the region falls during late winter to early summer, except in the eastern parts, where interannual variation is lower and the peak of precipitation is skewed towards the summer months. The decreasing precipitation pattern in the mountains from west to the east emanates the westerlies' dominance in moisture transport to Central

Asia (Bothe *et al* 2012, Jiang *et al* 2020, Peng *et al* 2021).

The moisture transport during February to June is modulated by multiple global ocean-atmospheric oscillations, some of which have longer temporal legacies with varying spatial effects on seasonal precipitation. Our findings suggest that the ENSO, PDO, EAWR, NAO, and SCAN are important determinants of precipitation intensity across Central Asia's mountains from February to June. This is consistent with earlier findings (e.g. Mariotti 2007, Wang *et al* 2014, Dixon and Wilby 2019, Gerlitz *et al* 2019), albeit our spatial scales are different and the targeted seasons do not precisely align.

ENSO and PDO exert spatially widespread and gradually evolving relationship with the peak precipitation season in the Central Asia mountains, that is detectable at extended lead times before the start of the precipitation season. As early as two to four months prior to the start of the season, the ENSO condition (as proxied by dominant lead time of SOI) and PDO are good predictors of seasonal precipitation across all subregions. In contrast, NAO and SCAN tend to have a shorter temporal span, emerging at the beginning of the season and are more pronounced at higher elevations. This partially restricts their value for forecasting. Reliable seasonal NAO and SCAN forecasts would hence improve seasonal precipitation forecasts, particularly for the southeastern part of the Central Asian mountains. The EAWR in October is another important predictor for the seasonal precipitation, which confirms other findings using different precipitation data and a target season from November to April (Gerlitz *et al* 2019).

The negative relationship between NAO and the seasonal precipitation contrasts prior findings. Earlier studies found a positive link between the NAO and cold season precipitation over Central Asia (Syed *et al* 2010, De Beurs *et al* 2018), whereas others reported a negative correlation (Hu *et al* 2017, Dixon and Wilby 2019, Zhang *et al* 2019). Our results suggest a negative correlation between the NAO and spring precipitation anomalies, which implies that NAO teleconnections are either more heterogenous across elevational gradients in Central Asia or non-stationary across seasonal scales.

Recent studies have reported on influence of positive winter SH anomalies on reduced annual precipitation totals in Central Asia (Zhang *et al* 2019), but these decreases likely occur mostly during winter (Riaz and Iqbal 2017). We have detected significant correlations between SH in December and the precipitation from February to June in the north eastern part of the Tian-Shan. A separate experiment (not shown here) suggests that inclusion of the SH in December as a predictor may contribute to the forecast accuracy for the Inner Tian-Shan. AMO and DMI exhibit stronger correlations one year ahead of the targeted season, but they fail to pass the field significance test across majority of the studied subregions. Similar patterns of delayed AMO linkages were found earlier for precipitation over northwest China (Zhong *et al* 2019). Additional research is necessary to ascertain whether AMO and DMI's long delayed association patterns with spring precipitation in Central mountains are random or display some mediating effect.

The SVR-based forecast models display robust performance, though they also tend to underestimate exceptionally rainy seasons. The prediction skill is better in those subregions where the interannual dynamics of precipitation arguably provide a stronger signal with less noise. In comparison, the flatter precipitation variations over the Eastern Pamir produce a poorer signal-to-noise ratio, resulting in lower forecast skill of SVR. Overall, the Eastern Pamir retains distinct patterns of the seasonal precipitation and its associations with most studied climate oscillations; they are either typically weaker in this subregion, shifted by several months, or even become inverse. Such inconsistent correlation patterns and subsequently the lower predictability of the seasonal precipitation for the Eastern Pamir could be also attributed to the comparatively weaker accuracy of the satellite-based precipitation estimates for this subregion (Peña-Guerrero *et al* 2022).

In sum, we highlighted the value of spatial and temporal aggregations for assessing climate teleconnections, which can be used to improve seasonal precipitation predictions. Our findings assert that appropriate climate regionalization helps to create spatial clusters with high internal homogeneity of precipitation so that their response to interannual variability could serve as a good target for seasonal prediction (Badr *et al* 2016) Climate teleconnections can also exhibit nonstationarity over time, which implies a need for careful definition of the target season; this is particularly relevant for mountainous regions where precipitation is more heterogeneous in space and time. Furthermore, we demonstrated how an understanding of the temporal dynamics of teleconnection can complement seasonal forecasting. Finally, machine learning-based seasonal precipitation forecasting demonstrated capable of handling the nonlinear and highly dimensional features that are inherent to climate teleconnections.

Data availability statement

The MSWEP v2.8 is available at www.gloh2o.org/mswep/. The oscillation indices were downloaded from online repositories of the Climate Research Unit https://crudata.uea.acuk/cru/data/soi/soi.dat (SOI), NOAA Climate Prediction Center www.cpc.ncep.noaa.gov/products/MD_index.php (NAO, PDO, AMO, DMI, SCAN, EAWR), Free University of Berlin www.geo.fu-berlin.de/met/ag/strat/produkte/qbo/qbo.dat (QBO).

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/tabumis/CA_hydroclimatic.

Acknowledgments

This research has been supported by the Volkswagen Foundation within the 'Structured doctoral programme on Sustainable Agricultural Development in Central Asia' (SUSADICA) project, Grant Number 96 264. We would like to thank the two anonymous reviewers and the editor whose valuable comments and suggestions helped to improve this manuscript. We are grateful to Rémi Cousin for his help with SEAS5 data retrieval and to Max Hofman for verifying the data preprocessing script for MSWX-Long data.

Code availability

The R script for correlation analysis and SVR modelling as well as the corresponding data (MSWEP v2.8 precipitation data for the study area, shapefile of the delineated subregions, and timeseries of the monthly oscillation indices) are available at https://github.com/tabumis/CA_hydroclimatic.

Conflict of interest

The authors declare that they have no conflict of interest.

ORCID iDs

Atabek Umirbekov https://orcid.org/0000-0001-5567-2574
Mayra Daniela Peña-Guerrero https://orcid.org/0000-0001-8974-2956
Daniel Müller https://orcid.org/0000-0001-8988-0718

References

Apel H, Abdykerimova Z, Agalhanova M, Baimaganbetov A, Gavrilenko N, Gerlitz L, Kalashnikova O, Unger-Shayesteh K, Vorogushyn S and Gafurov A 2018 Statistical forecast of seasonal discharge in Central Asia using observational records: development of a generic linear modelling tool for operational water resource management *Hydrol. Earth Syst. Sci.* **22** 2225–54

Badr H S, Dezfuli A K, Zaitchik B F and Peters-Lidard C D 2016 Regionalizing Africa: patterns of precipitation variability in observations and global climate models *J. Clim.* **29** 9027–43

Barnston A G and Livezey R E 1987 Classification, seasonality and persistence of low-frequency atmospheric circulation patterns *Mon. Weather Rev.* **115** 1083–126

Beck H E, Wood E F, Pan M, Fisher C K, Miralles D G, Van Dijk A I J M, McVicar T R and Adler R F 2019 MSWep v2 Global 3-hourly 0.1° precipitation: methodology and quantitative assessment *Bull. Am. Meteorol. Soc.* **100** 473–500

Benjamini Y and Hochberg Y 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing *J. R. Stat. Soc. Ser.* B **57** 289–300

Bothe O, Fraedrich K and Zhu X 2012 Precipitation climate of Central Asia and the large-scale atmospheric circulation *Theor. Appl. Climatol.* **108** 345–54

Brönnimann S, Malik A, Stickler A, Wegmann M, Raible C C, Muthers S, Anet J, Rozanov E and Schmutz W 2016 Multidecadal variations of the effects of the quasi-Biennial oscillation on the climate system *Atmos. Chem. Phys.* **16** 15529–43

Chantry M, Christensen H, Dueben P and Palmer T 2021 Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI *Phil. Trans. R. Soc.* A **379**

Chen X, Wang S, Hu Z, Zhou Q and Hu Q 2018 Spatiotemporal characteristics of seasonal precipitation and their relationships with ENSO in Central Asia during 1901–2013 *J. Geogr. Sci.* **28** 1341–68

Cohen J, Saito K and Entekhabi D 2001 The role of the Siberian High in Northern Hemisphere climate variability *Geophys. Res. Lett.* **28** 299–302

De Beurs K M, Henebry G M, Owsley B C and Sokolik I N 2018 Large scale climate oscillation impacts on temperature, precipitation and land surface phenology in Central Asia *Environ. Res. Lett.* **13** 065018

Dixon S G and Wilby R L 2019 A seasonal forecasting procedure for reservoir inflows in Central Asia *River Res. Appl.* **35** 1141–54

Enfield D B, Mestas-Nuñez A M and Trimble P J 2001 The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.* **28** 2077–80

Gerlitz L, Steirou E, Schneider C, Moron V, Vorogushyn S and Merz B 2019 Variability of the cold season climate in Central Asia. Part II: hydroclimatic predictability *J. Clim.* **32** 6015–33

Gerlitz L, Vorogushyn S and Gafurov A 2020 Climate informed seasonal forecast of water availability in Central Asia: state-of-the-art and decision making context *Water Secur.* **10** 100061

Gibson P B, Chapman W E, Altinok A, Delle Monache L, DeFlorio M J and Waliser D E 2021 Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts *Commun. Earth Environ.* **2** 159

GloH2O 2021 *MSWEP V2.8 Technical Documentation* pp 1–8 (available at: www.dropbox.com/s/5r4nnicfe3ft12d/MSWEP_V2_doc.pdf?dl=1)

Guan X, Yao J and Schneider C 2022 Variability of the precipitation over the Tianshan Mountains, Central Asia. Part II: multi-decadal precipitation trends and their association with atmospheric circulation in both the winter and summer seasons *Int. J. Climatol.* **42** 139–56

Hersbach H *et al* 2020 The ERA5 global reanalysis *Q. J. R. Meteorol. Soc.* **146** 1999–2049

Hu Z, Zhou Q, Chen X, Qian C, Wang S and Li J 2017 Variations and changes of annual precipitation in Central Asia over the last century *Int. J. Climatol.* **37** 157–70

Huffman G, Bolvin D, Braithwaite D, Hsu K, Joyce R, Kidd C, Nelkin E J, Sorooshian S, Tan J and Xie P 2019 Algorithm theoretical basis document version 6: integrated multi-satellite retrievals for GPM (IMERG) NASA (available at: https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf)

Immerzeel W W *et al* 2020 Importance and vulnerability of the world's water towers *Nature* **577** 364–9

Jiang J, Zhou T, Wang H, Qian Y, Noone D and Man W 2020 Tracking moisture sources of precipitation over central asia: a study based on the water-source-tagging method *J. Clim.* **33** 10339–55

Johnson S J *et al* 2019 SEAS5: the new ECMWF seasonal forecast system *Geosci. Model Dev.* **12** 1087–117

Karthe D, Abdullaev I, Boldgiv B, Borchardt D, Chalov S, Jarsjö J, Li L and Nittrouer J A 2017 Water in Central Asia: an integrated assessment for science-based management *Environ. Earth Sci.* **76** 690

Mantua N J, Hare S R, Zhang Y, Wallace J M and Francis R C 1997 A Pacific interdecadal climate oscillation with impacts on salmon production *Bull. Am. Meteorol. Soc.* **78** 1069–79
Mariotti A 2007 How ENSO impacts precipitation in southwest central Asia *Geophys. Res. Lett.* **34** 2–6
Naujokat B 1986 An update of the observed quasi-Biennial oscillation of the stratospheric winds over the tropics *J. Atmos. Sci.* **43** 1873–7
NRC 2010 *Assessment of Intraseasonal to Interannual Climate Prediction and Predictability* (Washington, DC: National Academies Press) (available at: www.nap.edu/catalog/12878)
Peña-Guerrero M D, Umirbekov A, Tarasova L and Müller D 2022 Comparing the performance of high-resolution global precipitation products across topographic and climatic gradients of Central Asia *Int. J. Climatol.* 1–16
Peng D, Zhou T and Zhang L 2021 Moisture sources associated with precipitation during dry and wet seasons over Central Asia *J. Clim.* **33** 10755–71
Portele T C, Lorenz C, Dibrani B, Laux P, Bliefernicht J and Kunstmann H 2021 Seasonal forecasts offer economic benefit for hydrological decision making in semi-arid regions *Sci. Rep.* **11** 1–15
R Core Team 2020 R: a language and environment for statistical computing (available at: www.r-project.org/)
Raghavendra S and Deka P C 2014 Support vector machine applications in the field of hydrology: a review *Appl. Soft Comput. J.* **19** 372–86
Riaz S M F and Iqbal M J 2017 Singular value decomposition analysis for examining the impact of Siberian High on winter precipitation variability over South Asia *Theor. Appl. Climatol.* **130** 1189–94
Ropelewski C F and Jones P D 1987 An extension of the Tahiti-Darwin Southern Oscillation Index *Mon. Weather Rev.* **115** 2161–5
Saji N H and Yamagata T 2003 Possible impacts of Indian Ocean Dipole mode events on global climate *Clim. Res.* **25** 151–69
Saji N, Goswami B, Vinayachandran P and Yamagata T 1999 A dipole mode in the Tropical Ocean *Nature* **401** 360–3
Satti S, Zaitchik B F, Badr H S and Tadesse T 2017 Enhancing dynamical seasonal predictions through objective regionalization *J. Appl. Meteorol. Climatol.* **56** 1431–42
Schär C, Vasilina L, Pertziger F and Dirren S 2004 Seasonal runoff forecasting using precipitation from meteorological data assimilation systems *J. Hydrometeorol.* **5** 959–73
Schiemann R 2007 Forcing and Variability of the Hydroclimate in Central Asia *Doctoral thesis* ETH (https://doi.org/10.3929/ethz-a-005629158)
Scholes R J 2020 The future of semi-arid regions: a weak fabric unravels *Climate* **8** 1–11
Syed F S, Giorgi F, Pal J S and Keay K 2010 Regional climate model simulation of winter climate over central-southwest Asia, with emphasis on NAO and ENSO effects *Int. J. Climatol.* **30** 220–35
Viviroli D and Weingartner R 2004 The hydrological significance of mountains: from regional to global scale *Hydrol. Earth Syst. Sci.* **8** 1017–30
Wang S, Huang J, He Y and Guan Y 2014 Combined effects of the Pacific Decadal Oscillation and El Niño-Southern Oscillation on Global Land Dry-Wet changes *Sci. Rep.* **4** 1–8
Ward J 1963 Hierarchical grouping to optimize an objective function *J. Am. Stat. Assoc.* **58** 236–44
Wilks D S 2006 On "field significance" and the false discovery rate *J. Appl. Meteorol. Climatol.* **45** 1181–9
Wilks D S 2016 The stippling shows statistically significant grid points *Bull. Am. Meteorol. Soc.* **97** 2263–74
World Bank 2018 *Central Asia Hydrometeorology Modernization Project (CAHMP) Additional Financing (P164780). Technical Project Paper* World Bank (available at: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/904441533353436310/central-asia-hydrometeorology-modernization-project-additional-financing)
Xenarios S, Gafurov A, Schmidt-Vogt D, Sehring J, Manandhar S, Hergarten C, Shigaeva J and Foggin M 2019 Climate change and adaptation of mountain societies in Central Asia: uncertainties, knowledge gaps, and data constraints *Reg. Environ. Change* **19** 1339–52
Zhang M, Chen Y, Shen Y and Li B 2019 Tracking climate change in Central Asia through temperature and precipitation extremes *J. Geogr. Sci.* **29** 3–28
Zhong F, Cheng Q and Ge Y 2019 Relationships between spatial and temporal variations in precipitation, climatic indices, and the normalized differential vegetation index in the upper and middle reaches of the Heihe River Basin, Northwest China *Water* **11** 1394